Uncorrectable Errors of Weight Half the Minimum Distance for Binary Linear Codes

Kenji Yasunaga[*]   Toru Fujiwara[+]

[*]Kwansei Gakuin University, Japan

[+]Osaka University, Japan

2008 IEEE International Symposium on Information Theory, July 6th to 11th, 2008
Sheraton Centre Toronto Hotel, Toronto, Ontario, Canada

Summary of the Work

Main Result

- A lower bound on #(uncorrectable errors of weight $\lceil d/2 \rceil$) for binary linear codes.
  - d : the minimum distance of the code
- A generalization to weight $> \lceil d/2 \rceil$.

Main Techniques

- Monotone error structure (Larger half)
  - Monotone error structure appears in [Peterson, Weldon, 1972] .
  - Larger half was introduced in [Helleseth, Kløve, Levenshtein, 2005] .

Outline

- Correctable/Uncorrectable Errors
- Our Results
- Monotone Error Structure
- Proof Sketch of Our Results

Outline

- Correctable/Uncorrectable Errors
- Our Results
- Monotone Error Structure
- Proof Sketch of Our Results

Problem Setting

- Binary linear code $C \subseteq \{0,1\}^n$
- Error vector $\boldsymbol{e} \in \{0,1\}^n$
- If $w(\boldsymbol{e}) < d/2 \Rightarrow \boldsymbol{e}$ is always correctable.
  If $w(\boldsymbol{e}) \geq d/2 \Rightarrow$?
  - $w(\boldsymbol{x})$: the Hamming weight of $\boldsymbol{x}$

In this work,

we investigate #(correctable errors of weight i) for $i \geq d/2$.

Correctable/Uncorrectable Errors

- Correctable errors $E^0(C)$
  = Correctable by minimum distance decoding.
  - $E_i^0(C)$: Correctable errors of weight i

- Uncorrectable errors $E^1(C) = \{0,1\}^n \setminus E^0(C)$
  - $E_i^1(C)$: Uncorrectable errors of weight i
  - $|E_i^0(C)| + |E_i^1(C)| = \binom{n}{i}$
  - The error probability over BSC$_p$ is $P_{error} = \sum_{i=0}^{n} p^i (1-p)^{n-i} |E_i^1(C)|$.

- Minimum distance decoding
  - Outputs a nearest (w.r.t. Hamming dist.) codeword to the input.
  - Performs ML decoding for BSC.
  - Syndrome decoding is a minimum distance decoding.

Syndrome Decoding

- Coset partitioning

$$\{0,1\}^n = \bigcup_{i=1}^{2^{n-k}} C_i, \quad C_i \cap C_j = \phi \ \text{ for } \ i \neq j$$

$$C_i = \{\boldsymbol{v}_i + \boldsymbol{c} : \boldsymbol{c} \in C\}$$: Coset of C

$$\boldsymbol{v}_i = \arg\min_{v \in C_i} w(\boldsymbol{v})$$: Coset leader of C_i

- Syndrome decoding
  - Output $\boldsymbol{y} + \boldsymbol{v}_i$ if $\boldsymbol{y} \in C_i$ ($\boldsymbol{y}$ is the input).
  - Coset leaders = Correctable errors.

Outline

- Correctable/Uncorrectable Errors
- Our Results
- Monotone Error Structure
- Proof Sketch of Our Results

Previous Results for $|E_i^1(C)|$

- For the first-order Reed-Muller code RM_m
  - $|E_{d/2}^1(\mathrm{RM}_m)|$ [Wu, 1998]
  - $|E_{d/2+1}^1(\mathrm{RM}_m)|$ [Yasunaga, Fujiwara, 2007]

- For binary linear codes
  - Upper bounds on $|E_i^1(C)|$ for every $0 \le i \le n$ [Poltyrev 1994], [Helleseth, Kløve 1997], [Helleseth, Kløve, Levenshtein 2005]

Our Results

- A lower bound on $|E^1_{\lceil d/2 \rceil}(C)|$ for codes satisfying some condition.
  - The condition is not too restrictive.
    - Long Reed-Muller codes and random linear codes satisfy
  - Given by #(codewords of weight d (and $d+1$)).
  - Asymptotically coincides with the corresponding upper bound for Reed-Muller codes and random linear codes.

- A generalization to $|E^1_i(C)|$ for $i > \lceil d/2 \rceil$.
  - The bound is weak.

Outline

- Correctable/Uncorrectable Errors
- Our Results
- Monotone Error Structure
- Proof Sketch of Our Results

Monotone Error Structure

- Recall that a coset leader is a minimum weight vector in a coset.

- There may be more than one minimum weight vector in the same coset.
  ⇒ Any of them will do.

- If we take the lexicographically smallest one for all cosets,
  ⇒ Correctable/uncorrectable errors have a monotone structure.

Monotone Error Structure

- Notation
  - Support of $\boldsymbol{v}$: $S(\boldsymbol{v}) = \{\ i : \boldsymbol{v}_i \neq 0\ \}$
  - $\boldsymbol{v}$ is covered by $\boldsymbol{u}$: $S(\boldsymbol{v}) \subseteq S(\boldsymbol{u})$

- Monotone error structure

  $\boldsymbol{v}$ is correctable.

  $\Rightarrow$ All vectors that are covered by $\boldsymbol{v}$ are correctable.

  $\boldsymbol{v}$ is uncorrectable.

  $\Rightarrow$ All vectors that cover $\boldsymbol{v}$ are uncorrectable.

- Example
  - 1100 is correctable. $\Rightarrow$ $0000, 1000, 0100$ are correctable.
  - 0011 is uncorrectable. $\Rightarrow$ $1011,\ 0111,\ 1111$ are uncorrectable.

Minimal Uncorrectable Errors

- Errors have the monotone structure (w.r.t $\subseteq$).
  $\Rightarrow$ $E^1(C)$ is characterized by minimal vectors (w.r.t. $\subseteq$).

- Minimal uncorrectable errors $M^1(C)$
  - = Uncorrectable errs. that are not covered by other uncorrectable errs.
  - $M^1(C)$ uniquely determines $E^1(C)$.

- Larger half $LH(\boldsymbol{c})$ of $\boldsymbol{c} \in C$
  - Introduced for characterizing $M^1(C)$ in [Helleseth et al., 2005].
  - Combinatorial construction is given in [Helleseth et al., 2005].
  - $M^1(C) \subseteq LH(C \setminus \{\mathbf{0}\}) \subseteq E^1(C)$, where $LH(S) = \bigcup_{c \in S} LH(\boldsymbol{c})$.

Outline

- Correctable/Uncorrectable Errors
- Previous Results
- Our Results
- Monotone Error Structure
- Proof Sketch of Our Results

Proof Sketch of Our Results

- Objective : To derive a lower bound on $|E^1_{\lceil d/2 \rceil}(C)|$.

- The following equalities hold:

$$M^1_{\lceil d/2 \rceil}(C) = LH_{\lceil d/2 \rceil}(C \setminus \{\mathbf{0}\}) = E^1_{\lceil d/2 \rceil}(C)$$

[Proof]

- $M^1(C) \subseteq LH(C \setminus \{\mathbf{0}\}) \subseteq E^1(C)$
- Since $\lceil d/2 \rceil$ is the smallest weight in $E^1(C)$, uncorrectable errors of weight $\lceil d/2 \rceil$ do not cover any other uncorrectable errors.

$\Rightarrow \quad M^1_{\lceil d/2 \rceil}(C) = E^1_{\lceil d/2 \rceil}(C)$

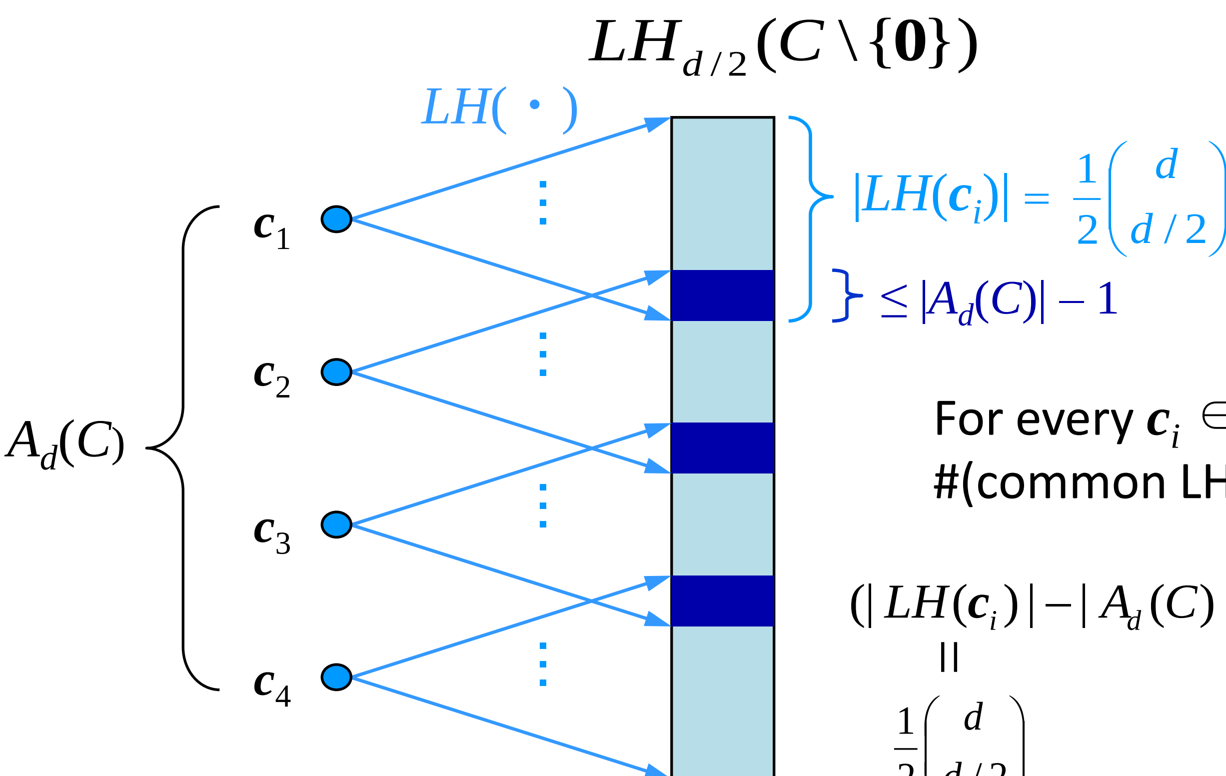
- Derive a lower bound on $|LH_{\lceil d/2 \rceil}(C \setminus \{\mathbf{0}\})|$.

Proof Sketch of Our Results (d is even)

- $LH_{d/2}(C \setminus \{\mathbf{0}\}) = LH(A_d(C))$, where $A_i(C)$ = { codewords of weight i in C}.

- Larger halves of two codewords in $A_d(C)$ are almost disjoint.

$$|LH(\boldsymbol{c}_1) \cap LH(\boldsymbol{c}_2)| \le 1 \quad \text{for every} \quad \boldsymbol{c}_1, \boldsymbol{c}_2 \in A_d(C)$$



For every $\boldsymbol{c}_i \in A_d(C)$,
#(common LH) is less than $|A_d(C)|$.

$$\underbrace{(|LH(\boldsymbol{c}_i)|}_{=\frac{1}{2}\binom{d}{d/2}} - |A_d(C)| + 1)\,|A_d(C)| \le \underbrace{|LH_{d/2}(C \setminus \{\mathbf{0}\})|}_{=|E^1_{d/2}(C)|}$$

The Results (d is even)

When d is even, if $\frac{1}{2}\binom{d}{d/2} > |A_d(C)| - 1$ holds, then

$$\frac{1}{2}\binom{d}{d/2}|A_d(C)| - (|A_d(C)|-1)|A_d(C)| \le |E^1_{d/2}(C)| \le \frac{1}{2}\binom{d}{d/2}|A_d(C)|.$$

Upper bound is from [Helleseth et al. 2005]

- If $|A_d(C)| \Big/ \binom{d}{d/2} \to 0$ as $n \to \infty$ then upper and lower bounds asymptotically coincide.
  - For Reed-Muller codes and random linear codes, the upper and lower bounds asymptotically coincide.

The Results (d is odd)

When d is odd, if $\binom{d}{(d+1)/2} > |A_d(C)| + |A_{d+1}(C)| - 1$ holds, then

$$\binom{d}{(d+1)/2}\left(|A_d(C)| + |A_{d+1}(C)|\right) - (2|A_d(C)| + |A_{d+1}(C)| - 1)|A_{d+1}(C)|$$

$$\leq \quad |E^1_{(d+1)/2}(C)| \quad \leq \quad \binom{d}{(d+1)/2}\left(|A_d(C)| + |A_{d+1}(C)|\right).$$

Upper bound is from [Helleseth et al. 2005]

- If $|A_{d+1}(C)| \Big/ \binom{d}{(d+1)/2} \to 0$ as $n \to \infty$ then upper and lower bounds asymptotically coincide.

A Generalization to Larger Weights

- A similar argument can be applied to weight $i > \lceil d/2 \rceil$.

For an integer i with $\lceil d/2 \rceil \le i \le \lfloor n/2 \rfloor$, if $\binom{2i-3}{i} > 3\binom{2i-\lceil d/2 \rceil}{i} B_i$ holds, then

$$\left(\binom{2i-3}{i} - 3\binom{2i-\lceil d/2 \rceil}{i} B_i \right) B_i \le |LH_i(C)| \le |E_i^1(C)|$$

$$\le \binom{2i-3}{i} |A_{2i-2}(C)| + 2\binom{2i-1}{i} (|A_{2i-1}(C)| + |A_{2i}(C)|)$$

where $B_i = |A_{2i-2}(C)| + |A_{2i-1}(C)| + |A_{2i}(C)|$.

For large i

- The condition for the bound is more restrictive.
- The bound is weak.
  - The bound is a lower bound on $LH_i(C)$.
  - The difference between $LH_i(C)$ and $E_i^1(C)$ is large.

Conclusion

Main results

- A lower bound on #(correctable errors of weight $\lceil d/2 \rceil$) for binary linear codes satisfying some condition.
  - The bound asymptotically coincides with the upper bound for Reed-Muller codes and random linear codes.
  - Monotone error structure & larger half are main tools.
  - A generalization to weight $i > \lceil d/2 \rceil$ is also obtained.
    - The generalized bound is weak for large i .

Future work

- A good lower bound for weight $> \lceil d/2 \rceil$.

Codes Satisfying the Condition

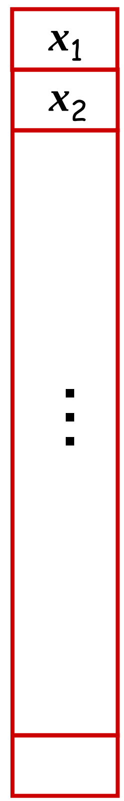
- The condition
$$\frac{1}{2}\binom{d}{d/2} > |A_d(T)| - 1 \quad \text{for even } d$$
$$\binom{d}{(d+1)/2} > |A_d(T)| + |A_{d+1}(T)| - 1 \quad \text{for odd } d$$

- Codes satisfying the condition
  - (n, k) primitive BCH codes for $n = 127$ and $k \leq 64$, $n = 63$ and $k \leq 24$
  - (n, k) extended primitive BCH codes for $n = 127$ and $k \leq 64$, $n = 63$ and $k \leq 24$
  - r-th order Reed-Muller codes of length 2^m fixed r and $m \to \infty$
  - Random linear codes for $n \to \infty$

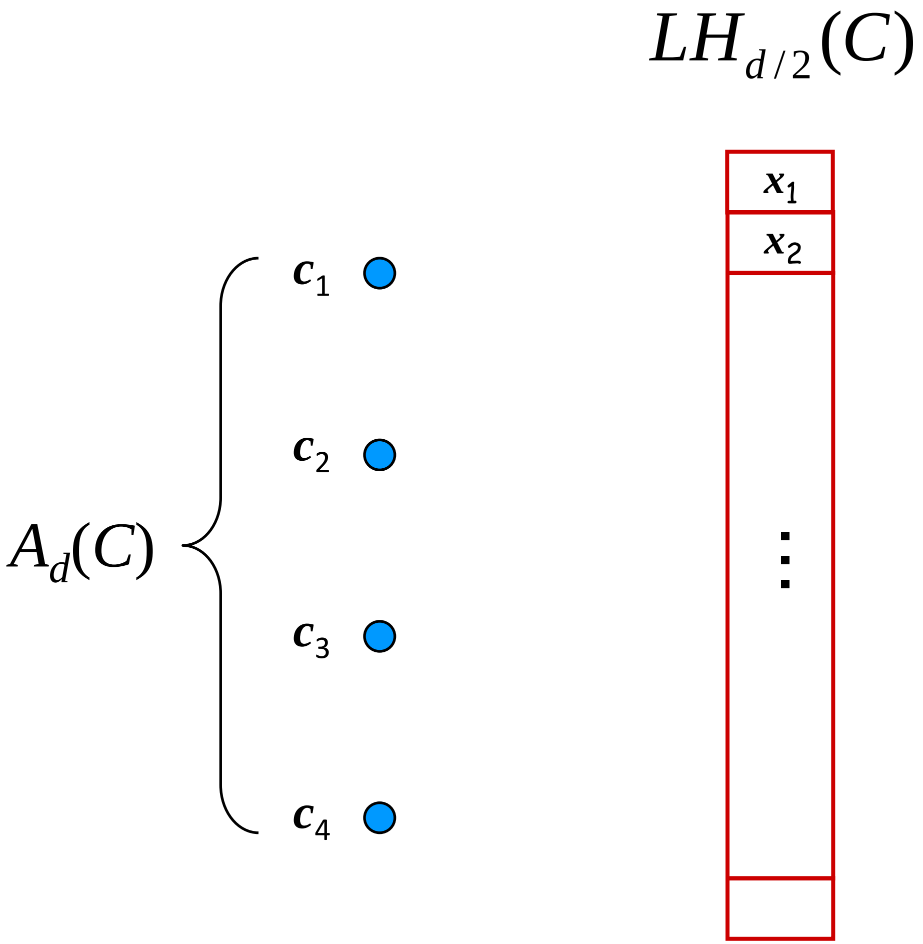
r	m
1	≥ 4
2	≥ 6
3	≥ 8
4	≥ 10
5	≥ 11
6	≥ 13

Proof Sketch of Our Results (d is even)

$$LH_{d/2}(C)$$

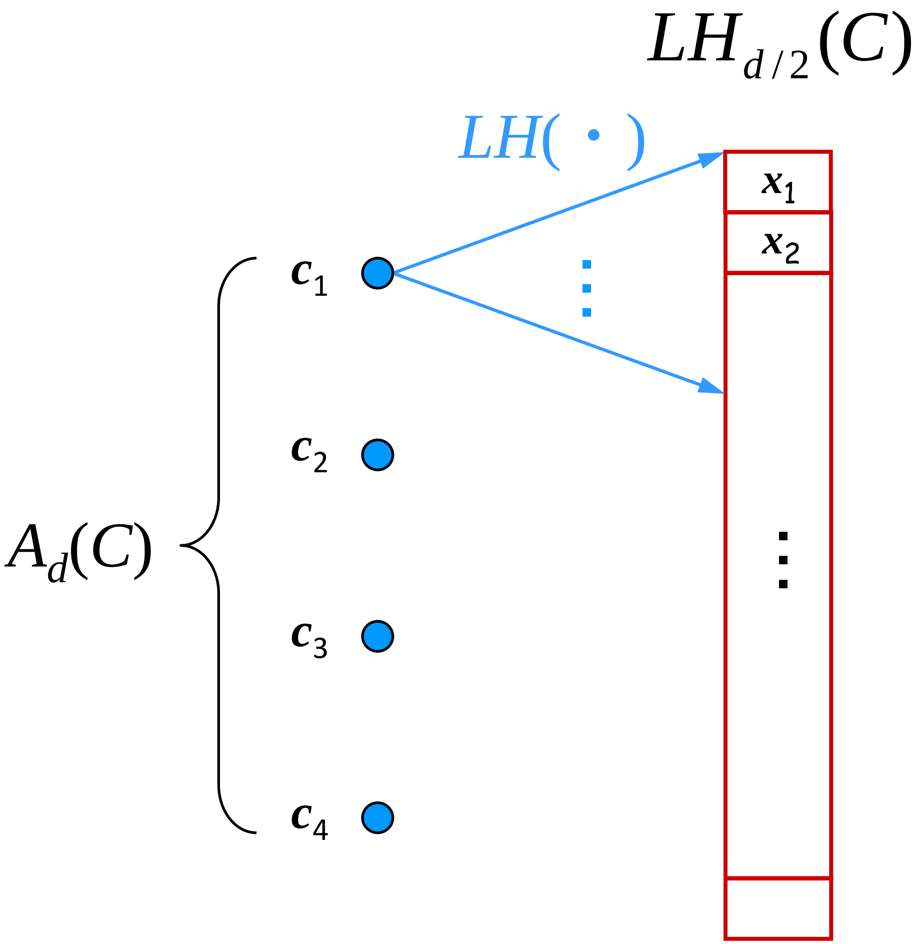


Proof Sketch of Our Results (d is even)



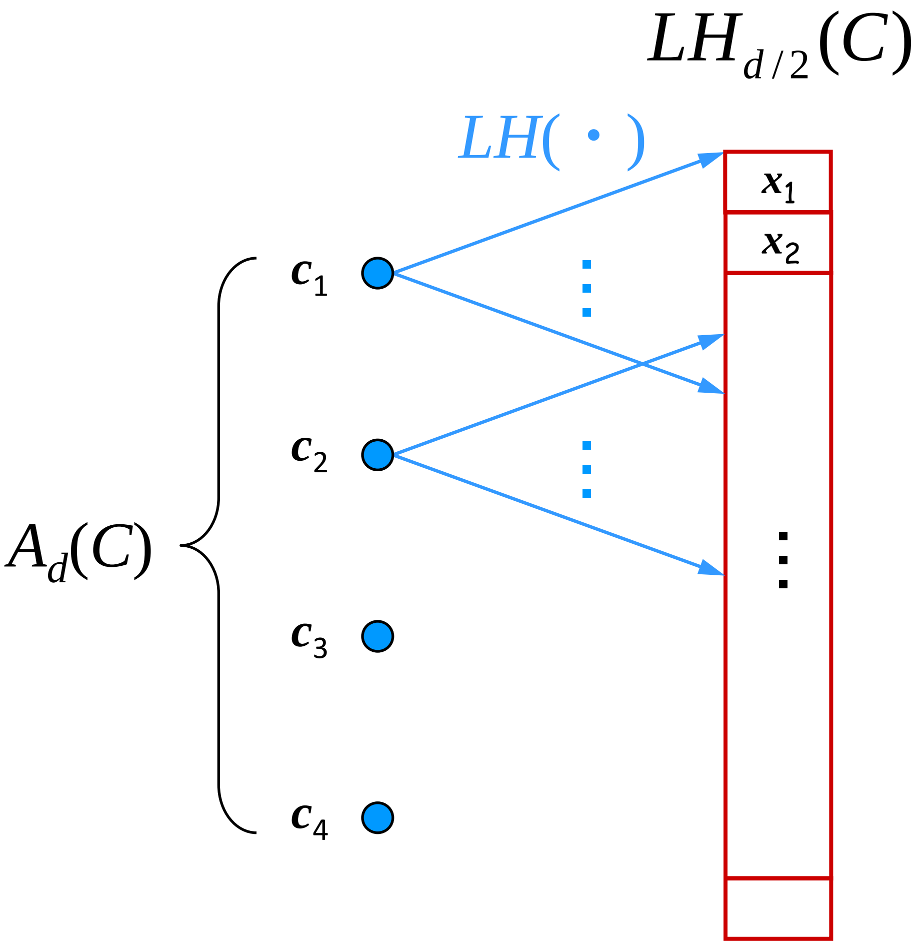
$A_i(C)$: the set of codewords with weight i

Proof Sketch of Our Results (d is even)



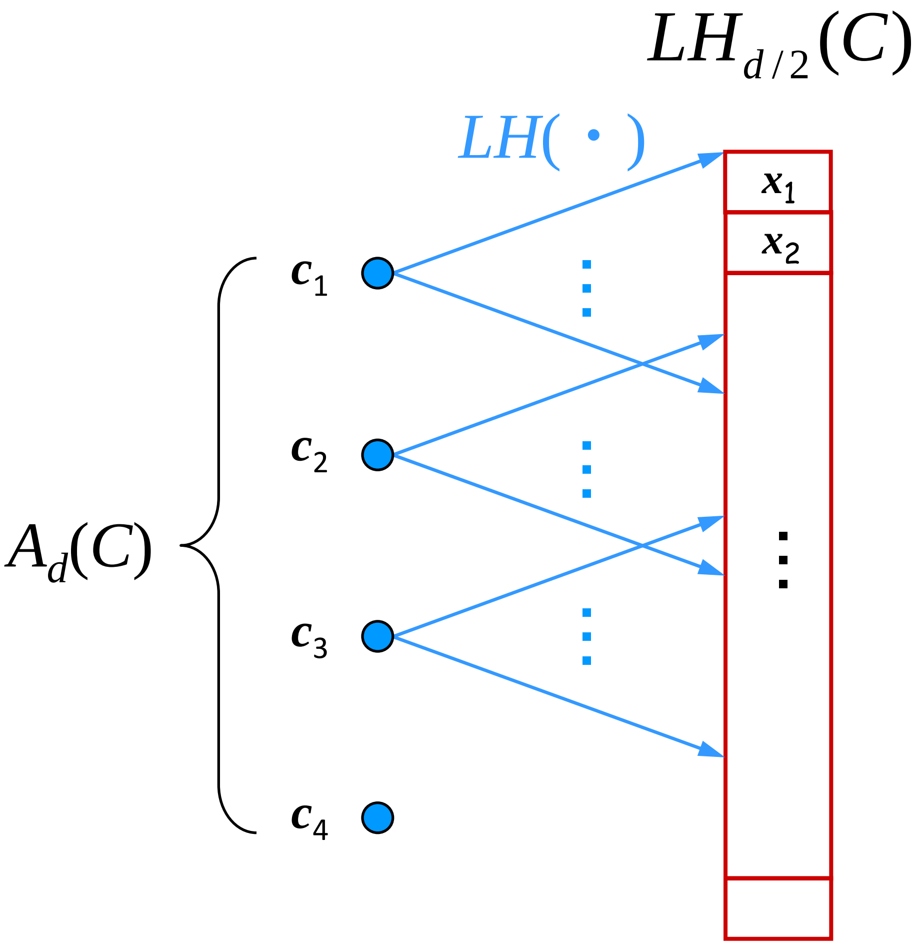
$A_i(C)$: the set of codewords with weight i

Proof Sketch of Our Results (d is even)



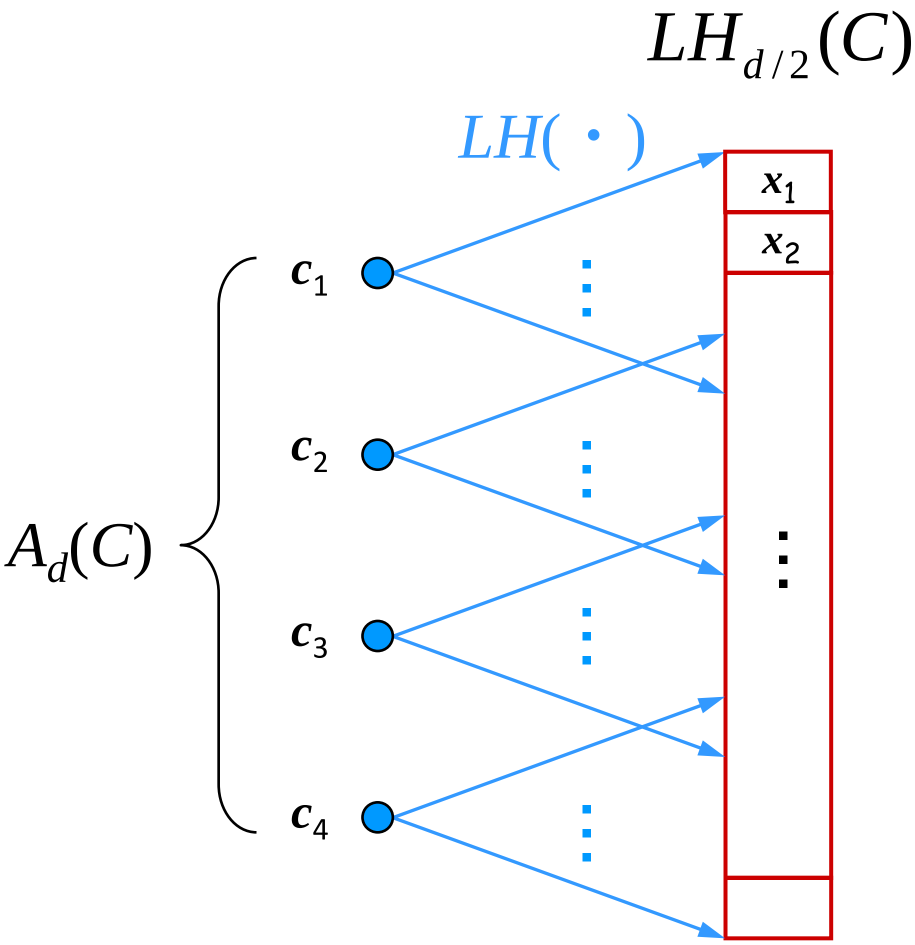
$A_i(C)$: the set of codewords with weight i

Proof Sketch of Our Results (d is even)



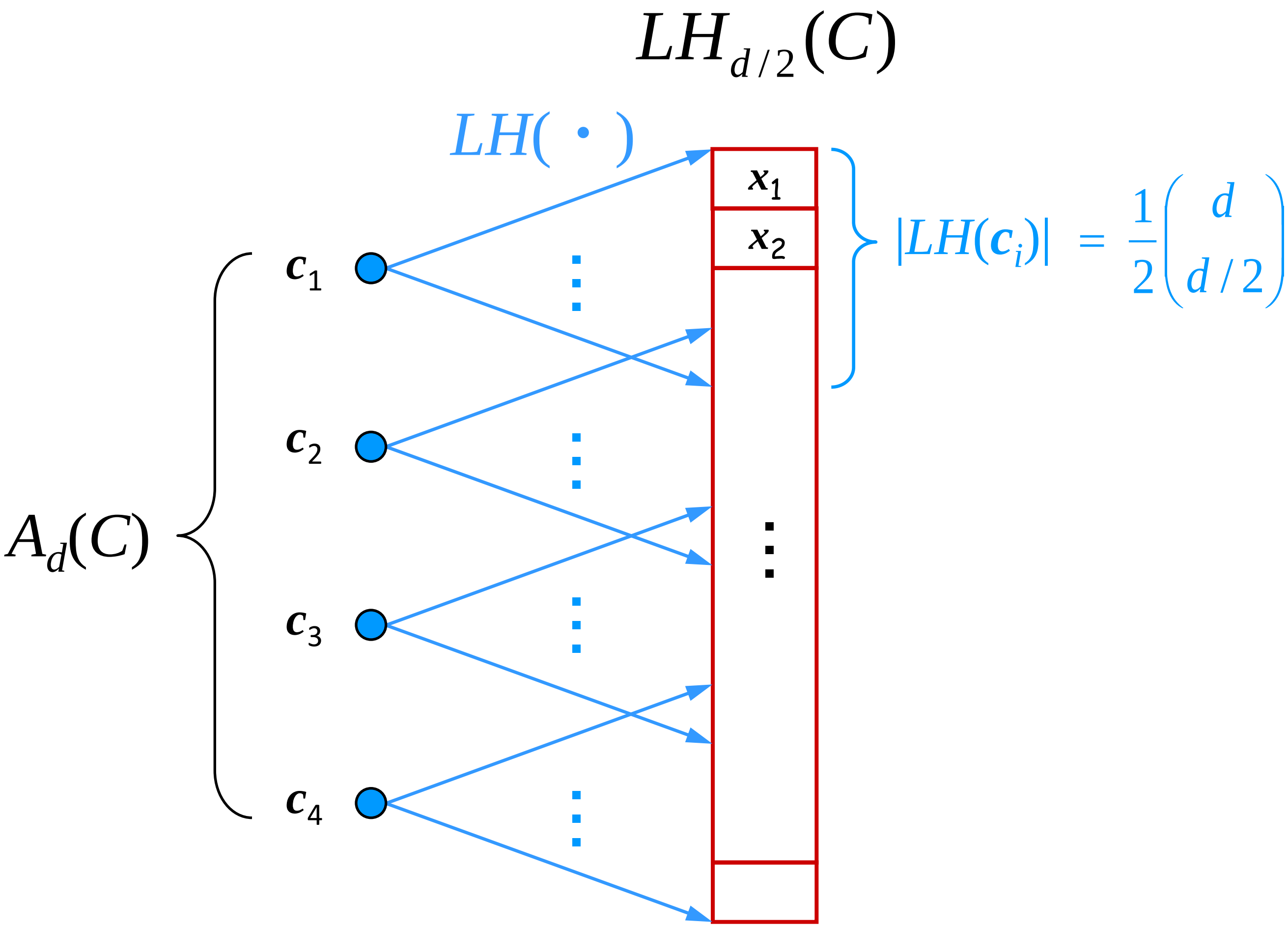
$A_i(C)$: the set of codewords with weight i

Proof Sketch of Our Results (d is even)



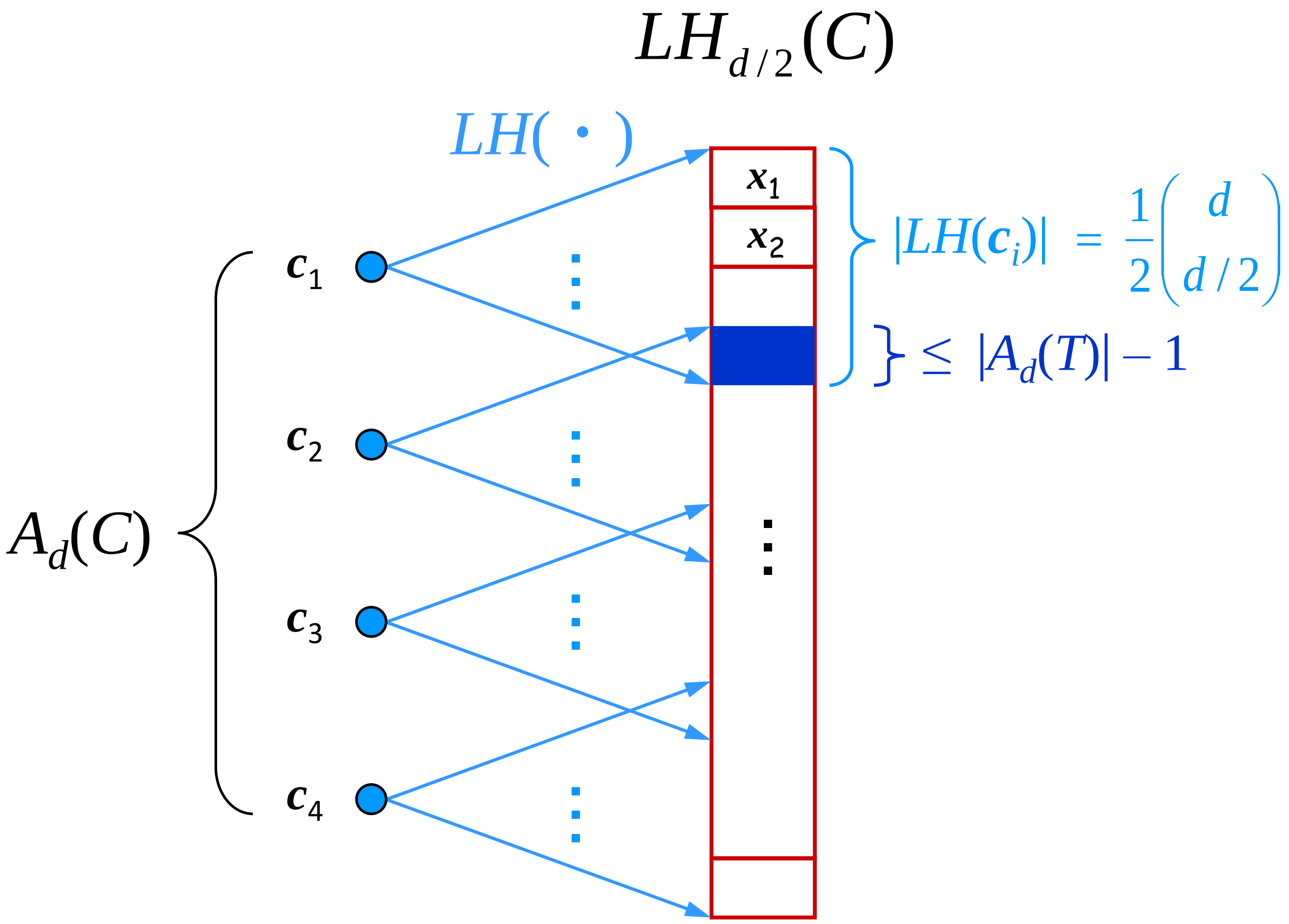
$A_i(C)$: the set of codewords with weight i

Proof Sketch of Our Results (d is even)



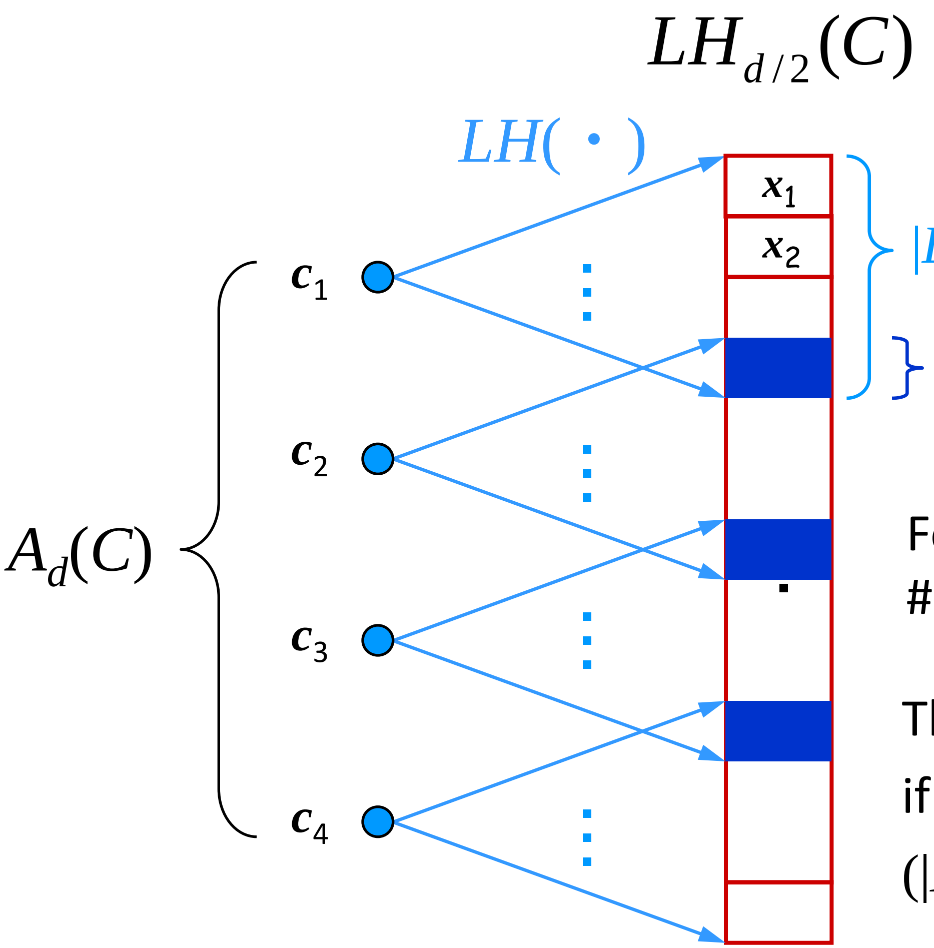
$A_i(C)$: the set of codewords with weight i

Proof Sketch of Our Results (d is even)



$A_i(C)$: the set of codewords with weight i

Proof Sketch of Our Results (d is even)



For every $\boldsymbol{c}_i \in A_d(T)$,
#(common LH) is less than $|A_d(T)| - 1$

Thus,

if $|LH(\boldsymbol{c}_i)| > |A_d(T)| - 1$, then

$(|LH(\boldsymbol{c}_i)| - |A_d(T)| + 1)\,|A_d(T)| \leq |LH_{d/2}(T)|$

$A_i(C)$: the set of codewords with weight i

A Generalization to Larger Weight

- The lower bound for weight $\lceil d/2 \rceil$ is obtained by considering the vectors of weight $\lceil d/2 \rceil$ in

$$M^1(C) \subseteq LH(C) \subseteq E^1(C)$$

- A similar argument can be applied to weight $i \geq \lceil d/2 \rceil + 1$ However, for large i,
  - The condition for the bound is more restrictive
  - The bound is weak
    - The bound is a lower bound on $LH_i(C)$
    - The difference between $LH_i(C)$ and $E_i^1(C)$ is large